IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 9 December 2024, accepted 23 December 2024, date of publication 3 January 2025, date of current version 8 January 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3525996

APPLIED RESEARCH

An Effective Deep Neural Network Architecture for EEG-Based Recognition of Emotions

KHADIDJA HENNI[1,2], NEILA MEZGHANI[1,2], AMAR MITICHE[3], LINA ABOU-ABBAS[1,2], AND AMEL BENAZZA-BEN YAHIA[4]

[1]Applied Intelligence Artificial Institute, TÉLUQ University, Montreal, QC G1K 9H6, Canada
[2]Imaging and Orthopaedics Research Laboratory, The CHUM Research Center, Montreal, QC H2X 0A9, Canada
[3]INRS—Centre Énergie, Matériaux et Télécommunications, Montréal, QC H5A 1K6, Canada
[4]COSIM Laboratory, SUP'COM, University of Carthage, Tunis 1054, Tunisia

Corresponding author: Khadidja Henni (khadidja.henni@teluq.ca)

This work was supported in part by Canada Research Chair on Biomedical Data Mining under Grant 950-231214.

ABSTRACT Emotions are caused by a human brain reaction to objective events. The purpose of this study is to investigate emotion identification by machine learning using electroencephalography (EEG) data. Current research in EEG-based emotion recognition faces significant challenges due to the high-dimensionality and variability of EEG signals, which complicate accurate classification. Traditional methods often struggle to extract relevant features from noisy and high-dimensional data, and they typically fail to capture the complex temporal dependencies within EEG signals. Recent progress in machine learning by deep neural networks has opened up opportunities to develop methods highly efficient and practicable as to serve useful real-world applications. The purpose of this study is to investigate a novel end-to-end deep learning method of emotion recognition using EEG data, which prefaces a combination of two-dimensional (2D) convolutional network (CNN) and Long short-term memory network (LSTM) by an autoencoder. The autoencoder layers seek a lower dimensionality encoding for optimal input signal reconstruction, and the 2D CNN/LSTM combination layers capture both spatial and temporal features that best describe the emotion classes present in the data. Experiments in four-category classification of emotions, using the public and freely available DEAP dataset, revealed that the method reached superior performance: 90.04% for the ‘‘arousal’’ category, 89.97% for ‘‘valence’’, 87.73% for ‘‘dominance,’’ and 90.84% for liking’’, as measured by the accuracy metric.

INDEX TERMS EEG signal, emotion recognition, auto-encoder, LSTM, CNN.

I. INTRODUCTION

This Emotions often occur in human interaction, conditioning subsequent behaviours and actions. As such, the ability of computers to recognize emotions can serve human-computer interaction (HCI) by providing subsequent other functions with information to guide processing for improved conclusions. This has strongly motivated recent research in HCI using emotion recognition, often referred to as affective computing, which seeks outcomes conditioned on subjects emotional states. As an essential component of affective computing, emotion recognition has, therefore, attracted great interest in research, and applications as well [1].

The associate editor coordinating the review of this manuscript and approving it for publication was Filbert Juwono.

This study investigates emotion recognition using electroencephalography (EEG) data, which records brain temporal electrical activity via electrodes positioned on the scalp. Visual examination of these records by experts is not practicable in most applications, including emotion recognition, because it is obviously substantially tedious and time consuming [2], [3]. There have been several studies of EEG-based pattern recognition, for epileptic seizure detection, and prediction, for instance [4], [5], [6]. Such tasks are confronted with the high variability and high dimensionality of the EEG recorded signals [7], have been effectively resolved by deep neural networks [6].

To address these challenges, deep neural networks [8] have proven effective in learning relevant features from raw EEG data and characterizing the EEG pattern classes of interest, and can do so in the face of high data dimensionally and high variability. Studies like [9] and [10] highlight the versatility of neural networks, further supporting their use in this study for EEG-based emotion recognition.

Consequently, this property can be determinant in successful emotion recognition from EEG data, and in establishing an explicit correspondence between EEG features and the emotions they characterize [11].

Effective emotion recognition using deep neural networks provides valuable insights into the fundamental mechanisms of emotive computing, showcasing their potential in supporting potent and useful applications [12].

Convolutional networks (CNN) are commonly used in deep machine learning [13]. They have been used as effective solvers in a variety of biomedical data classification problems, including EEG-based emotion recognition [11]. From a broad perpective, their popularity can be attributed to their exceptional results in several useful applications, and to their simple conceptual interpretation as networks of successive signal filtering layers. However, conceived primarily for images, this filtering layer design does not naturally accommodate time series data, such as EEG data of interest in this study. Long short-term memory networks (LSTM) [14] are recurrent neural networks (RNN) [15], [16] which correct this shortcoming by learning time dependencies to determine temporal features which can be crucial in time series prediction problems, such as EEG data classification. Further improvements have been obtained by CNN/LSTM network combination to emphasize spatial/temporal feature learning [1]

While CNN-LSTM architectures have demonstrated their efficacy in combining spatial and temporal feature learning, challenges persist in handling the high dimensionality and variability of EEG data. This motivates the integration of additional processing layers to further enhance feature extraction and classification accuracy.

This study introduces a novel framework that integrates an autoencoder layer for dimensionality reduction and relevant feature presentation, 2D convolutional layers for enhanced spatial feature extraction, and an optimized LSTM structure for capturing temporal dependencies. This combination addresses key challenges such as high dimensionality and variability in EEG signals, providing a robust solution for emotion recognition. The autoencoder layers are specifically designed to reduce dimensionality while retaining the most relevant features of the input signal, effectively addressing the variability and complexity inherent in EEG data. This step is particularly crucial for EEG signals, which are inherently noisy and high-dimensional, as it simplifies the data representation while preserving essential patterns necessary for accurate emotion recognition. Following this process, the 2D-CNN layers extract spatial features, and the LSTM layers capture temporal dependencies, creating a synergistic framework for spatiotemporal feature learning. This architecture combines the strengths of dimensionality reduction, spatial representation, and temporal modeling to achieve robust emotion classification performance.

As described in detail subsequently, the proposed method has been tested on raw data of the DEAP dataset, which consists of EEG signals collected from 32 participants exposed to 40 one-minute videos. The signals have been preprocessed, and each entry in the dataset (40960*8064) was associated with four binary labels representing the recent emotion as "arousal," "valence," "dominance," and "liking." The proposed approach demonstrated exceptional performance in 4-category emotion classification, achieving accuracy rates of 90.04% for "arousal," 89.97% for "valence," 87.73% for "dominance," and 90.84% for "liking" categories, underscoring its efficacy in capturing emotional nuances.

By addressing critical challenges such as high dimensionality and variability, these contributions advance the state-of-the-art in EEG-based emotion recognition. The proposed framework achieves superior classification accuracy, underscoring its potential to improve real-world applications in emotion recognition systems.

The remainder of this paper is organized as follows: Section II (Related studies) positions this study in the context of others, Section III (Material and methods) describes the data, the dataset, the basic deep-learning paradigms used, and the proposed method in detail. Section IV (Experimental results) presents the proposed method evaluation, as well as a discussion of the results. Finally, Section V contains a conclusion.

II. RELATED STUDIES

Existing EEG analysis methods for emotion recognition can be categorized into either shallow or deep learning approaches, both of which may involve signal preprocessing and feature extraction techniques.

Feature extraction is, of course at the heart of pattern classification. From a broad perspective, one can distinguish three categories of pattern features, namely, time domain characteristics, frequency domain, and joint time and frequency domain. For EEG data analysis, time domain features have targeted the signal temporal variation information, such as higher-order crossing [2]. Frequency domain features have been sought to characterize the EEG signal by its frequency contents, for instance using the Fast Fourier transform (FFT) [3], [11]. Along this vein of description, the Short-time Fourier transform (STFT) is often used to determine features which capture the signal temporal variation information [12]. The wavelet transform has been an alternative to the Fourier transform to allow better spatial localization of the information [14], [15], [23], [25].

A variety of emotion descriptive features have been constructed from spatial or frequency information, for instance based on differential entropy (DE) [17], fractal representation [18], power spectral density (PSD) [19], Deep forests [20], empirical mode decomposition (EMD) [21], [22] and discrete wavelet transform [23], [25]".

Classification is just as important as feature extraction. Generally, deep neural networks are distinguished from other classifiers, called shallow classifiers. Several common shallow classifiers have been considered in EEG-based emotion classification [26], for instance support vector machines (SVM), and variants such as the linear SVM [18], adaptive SVM [27], polynomial SVM [28], and RBF SVM [29], as well as the K-Nearest Neighbor (KNN) [30], and fuzzy clustering [31]. The accuracy of shallow classifiers is usually dependant upon, and sensitive to, the effectiveness of feature extraction. Deep learning neural networks do correct this shortcoming of shallow classifiers, by combining feature extraction and classification in their computational architecture. Several of such networks have been investigated for EEG-based emotion recognition [31], [34], [35], [36], [37], [38], [39], [41]. The evaluation of deep learning networks for EEG-based emotion recognition is usually tested on dedicated datasets such as SEED [32] and DEAP [33].

We mention here the main characteristic of the architecture of these deep neural networks. The study [35] describes a network, called SincNet, composed of three CNN layers connected to three DNN neural network layers, with batch normalization between the layers. In [36], the network extracts wavelet-based properties as features; it uses an architecture consisting of an LSTM, followed by fully connected layers and a prediction layer. This network reached good performance on the SEED dataset.

The network in [34] combined an SVM with two CNNs, LeNet and ResNet, with moderate success on the SEED database.

Other deep networks architectures for emotion recognition, which used the DEAP dataset, include [37], which extracts temporal frequency information by combining CNN and LSTM layers, with SoftMax prediction, [38] with CNN/DNN architecture, [39] with LSTM layers and a prediction layer, [31] with a Stack Auto-Encoder SAE and LSTM-RNN layers, and [40] with an CNN-LSTM combination.

Recent studies have also explored innovative approaches for emotion recognition using deep learning. In [51], Topic and Russo developed an emotion recognition model based on EEG feature maps through a deep learning network, achieving significant performance improvements. Another study by Li et al. [52] proposed a novel ensemble learning method using multiple objective particle swarm optimization for subject-independent EEG-based emotion recognition, demonstrating robustness across different subjects. Additionally, Akhand et al. [53] enhanced EEG-based emotion recognition by improving connectivity feature maps, providing insights into the intricate relationships between different brain regions during emotional responses.

Evaluation of emotion recognition has generally used either the SEED or the DEAP database, but most studies have favoured the DEAP data set because of its relatively large size. SEED is composed of data collected from 15 participants, featuring EEG data with an average data collection duration of approximately 15 minutes, focusing on negative, neutral, and positive emotions. In contrast, DEAP is composed of data collected from 32 participants who watched 40 one-minute music video excerpts, rating arousal, valence, liking/disliking, dominance, and familiarity. DEAP also includes physiological signals. While SEED emphasizes social interactions, DEAP aims to induce emotions through controlled stimuli, offering a broader range of evaluated emotions.

III. MATERIALS AND METHODS

A. EMOTIONS AND EEG ANALYSIS

Emotions are complex psychological and physiological states typically associated with feelings of pleasure or displeasure which cause related action. Emotion are often represented by the dimensional model [42], and encompass four key dimensions. Valence represents the positivity or negativity of an emotion, arousal measures its intensity, dominance indicates the sense of control, and liking reflects personal preference or enjoyment. These dimensions collectively provide a comprehensive understanding of the complex psychological and physiological states associated with emotions. EEG signals typically vary according to emotions, and studies [43] have shown explicitly this relationship, so that research has sought computer-assisted methods to automatically detect and classify emotions from EEG readings.

EEG data acquisition consists of placing several electrodes on a subject scalp to measure and record the brain electrical activity. The electrodes are placed according to the standard 10/20 international electrode placement scheme [44], where adjacent electrodes are separated by 10% or 20% of the distance between the front and back of the skull, or between the left and right of the skull. The EEG signal is recorded by measuring the voltage difference between a scalp-placed electrode and a reference electrode (mono-polar recording) or by measuring the voltage difference between two scalp-placed electrodes (bi-polar recording). The recorded amplitude of EEG is generally in the interval [10,100] mV and the frequency in [1,100] Hz. The EEG signal may be classified into five frequency bands: Delta band [0,4] Hz, Theta [4,8] Hz, (3) Alpha [8,14] Hz, (4) Beta [14,40] Hz, and Gamma band [40,100] Hz.

B. DEAP DATASET

Several datasets are available for research in emotion recognition, such as DEAP (Dataset for Emotion Analysis using Physiological Signals) [33], SEED (The SJTU Emotion EEG Dataset) [32], and MAHNOB (The MAHNOB Laughter Database) [45].

This study uses the DEAP database to carry out experiments due to its larger dataset, coverage of various emotions, and inclusion of diverse physiological signals. It contains EEG, and also ECG, EMG, plethysmographs, temperature, breathing zone, and other physiological signals. It also has videos of 32 participants watching 40 videos of 1 min to stimulate different emotions. The subjects are 16 males and 16 females, with ages between 19 and 37. They wore 40 electrodes and evaluated their emotional responses to the 40 videos across four dimensions: valence (indicating positivity/negativity), arousal (indicating excitement), dominance (indicating control), and liking (indicating their familiarity). These assessments were made using a scale from 1 to 9 with the self-assessment manikin [46].

The recorded data has a duration of 63 seconds, with the initial 3 seconds serving as a pretrial baseline, followed by 60 seconds of trailing signals. These signals have undergone preprocessing, including the application of a band-pass frequency filter and down-sampling to 128 Hz to enhance data quality. For classification purposes, the four labels are considered: valence, arousal, dominance and liking. However, to facilitate a binary classification task, a threshold is applied to partition the scale. Values greater than 5 represent positive emotions, while values less than or equal to 5 signify negative emotions, as outlined in previous studies [47], [48]. This approach simplifies the classification process and enables a more focused analysis of emotional states.

Our method has been tested on raw data of the DEAP dataset, which was structured into a learning matrix of 40960 entries and 8064 dimensions. Entries are the segments calculated from 32 participants, each watching 40 videos, with 32 channels (40960 = 32 * 40 * 32). The dimension of the dataset is 8064 which represents the entire EEG signal, such as segments of data points over 63 seconds and 128 Hz sampling, leading to the processed form of 8064 dimensions (8064=63*128). We consider all channels from the same EEG for each subject and video as inputs, since it is uncertain which channels will best capture the reactions due to the varying responses from different regions of the brain. Each entry was associated with four binary labels representing the emotional states: "arousal", "valence", "dominance" and "liking".

This rich dataset offers an extensive range of emotions and diverse physiological signals, making it an ideal choice for our experiments. The DEAP dataset has been widely used in research on emotion recognition to evaluate a variety of different methods, including traditional and deep neural classifiers. Overall, the DEAP dataset is a valuable resource for researchers studying emotion recognition and the use of physiological signals in emotion research.

C. PROPOSED METHOD

The EEG signals have been denoised and all artifacts were removed by a low pass filter with cutoff frequency equal to 50 Hz. The role of this filter is to remove frequencies higher than 50 Hz, which mainly correspond to noise. The figure 1 represents a filtered EEG channel.

Our emotion recognition model, as illustrated in Figure 2, leverages the power of deep learning. It consists of two integral components. Initially, an autoencoder is employed to extract salient features from EEG signals. Subsequently, the classification is done by the hybrid architecture that seamlessly integrates CNN and LSTM networks. This combined approach enables the model to effectively decipher emotions embedded in EEG data.

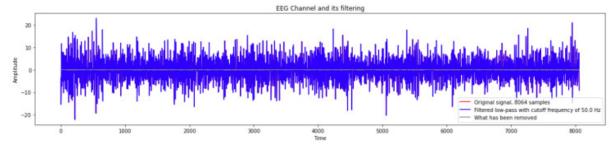


FIGURE 1. Filtered EEG channel by the low pass-filter with cutoff frequency = 50Hz.

The framework operates through three main stages, each addressing specific challenges of EEG data processing:

1) Dimensionality Reduction: The autoencoder processes the high-dimensional EEG signals to reduce their complexity while preserving the most relevant features. This step enhances the efficiency of subsequent layers by focusing on essential patterns in the data.
2) Spatial Feature Extraction: The 2D convolutional layers analyze the processed signals to identify critical spatial patterns. These layers are particularly effective in capturing local structures crucial for emotion classification.
3) Temporal Feature Learning: The optimized LSTM layers focus on capturing temporal dependencies, allowing the framework to model sequential variations in the EEG signals that correspond to emotional states.

Each of these stages is explicitly represented in Figure 2, ensuring a clear and intuitive understanding of the processing flow.

1) AUTOENCODER FOR FEATURE EXTRACTION

The Auto-Encoder is a potential extractor of spatially local relevant data features, but it is less efficient to capture long-term dependence relationships in sequence data, which can be easily fixed by LSTM [40]. The hybrid models have shown good performances in signal analysis. Therefore, AutoEncoder is used to extract features from raw EEG signals and then these features are the input of the hybrid network model CNN-LSTM. The Auto-Encoder is used for feature engineering and the CNN-LSTM network followed by prediction layers are used for the classification of raw EEG signals. The architecture of our deep Auto-Encoder is presented in Figure 3, it consists of an input layer with 512 dense nodes, fully connected to the first hidden layer with 256 dense nodes. This is followed by a bottleneck layer with 64 dense nodes, crucial for compressing the data, containing 64 dense nodes. This compressed representation, or latent space, captures the most significant features of the data. The decoder mirrors the encoder with a hidden layer of 256 dense nodes and an output layer of 512 dense nodes, where the reconstructed format of input data is generated.

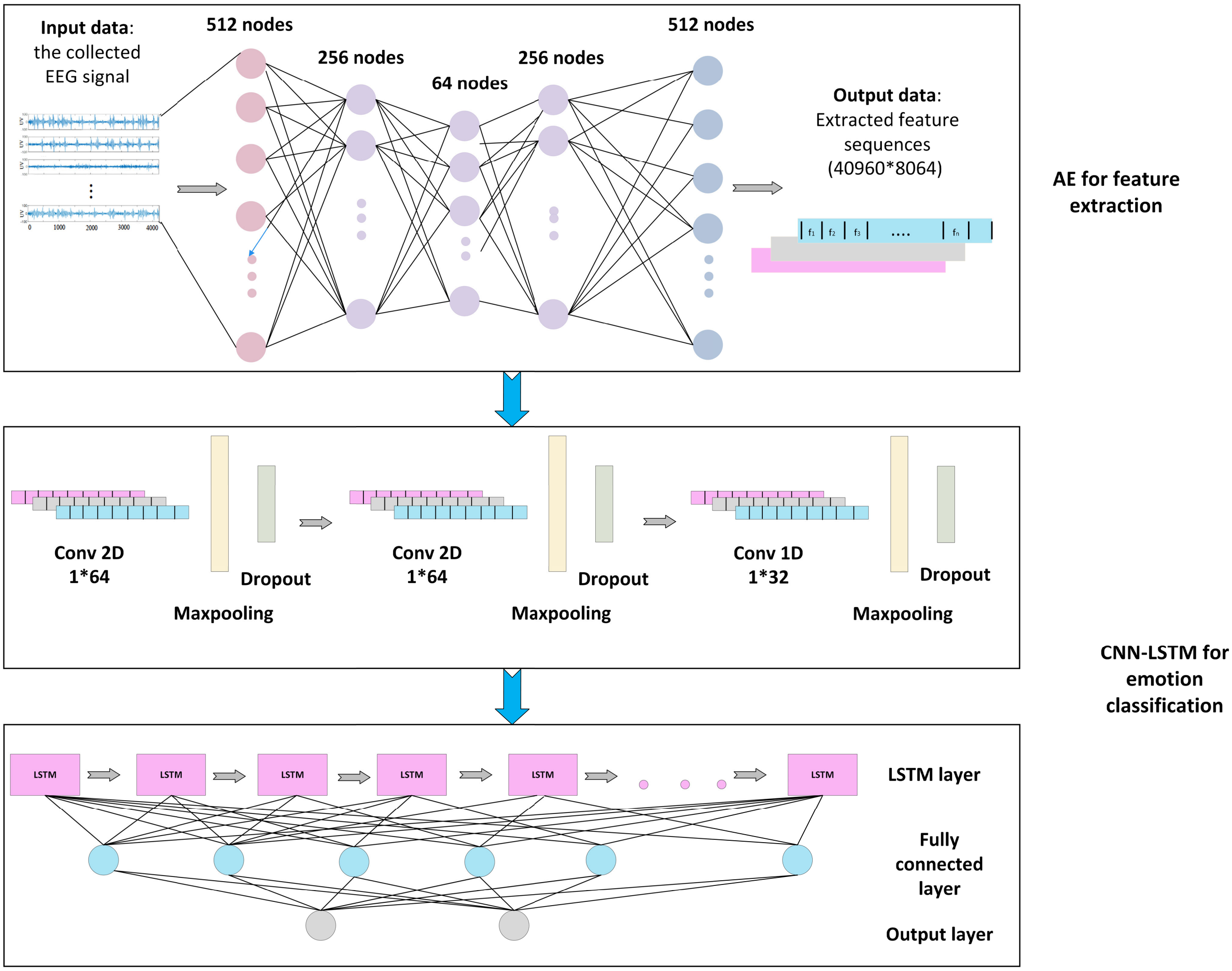


FIGURE 2. The proposed architecture auto-encoder+ CNN-LSTM model.

This structure allows for efficient feature extraction by distilling the essential characteristics of the EEG signals into a smaller dimensional space, 8064 features per sample are reduced to a more manageable 512 feature, thereby facilitating more effective classification in subsequent steps.

2) CNN-LSTM MODEL FOR EMOTION CLASSIFICATION

Convolutional neuronal networks, CNNs are a deep learning architectures tailored for processing grid-like data, such as images or, in your case, EEG feature maps. CNNs specialize in capturing spatial relationships within data. Their fundamental components include convolutional layers, which apply filters or kernels to input data, enabling them to learn and extract spatial features effectively.== After convolution, max pooling layers are often applied to downsample feature maps by selecting the maximum values from local regions. Dropout layers are employed for regularization, randomly deactivating a fraction of neurons during training to prevent overfitting.

Long Short-Term Memory, commonly referred to as LSTMs, are a specialized type of recurrent neural network (RNN). They are designed to capture and understand temporal dependencies within sequential data. LSTMs are particularly well-suited for tasks involving data that evolves over time, making them a valuable tool in EEG signal analysis. The core components of LSTM networks include LSTM cells, which contain various gates (input gate, output gate, forget gate, and cell gate) that regulate the flow of information within the network. These gates enable LSTMs to recognize and remember patterns in data sequences. LSTMs employ a training technique known as Backpropagation Through Time (BPTT), which allows them to learn and update their internal states over time, making them adept at modeling temporal dynamics.

In our model, the input data, after passing through the Auto-Encoder, the dimension of the 40960 samples are

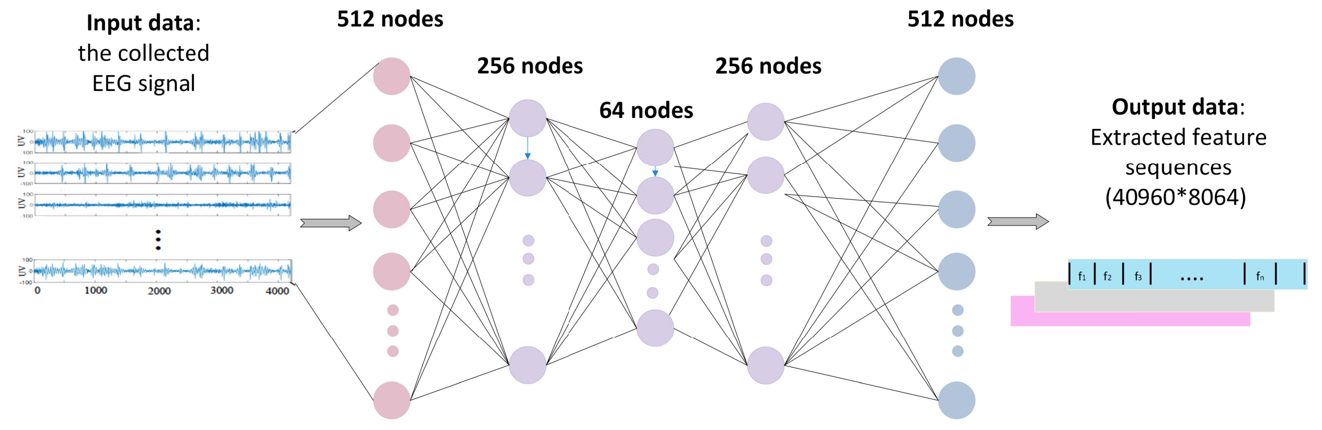


FIGURE 3. The proposed auto-encoder architecture.

reduces each with 64 features. These are then transformed into feature maps through three convolutional layers. Our CNN-LSTM network integrates three convolutional layers, each using different filter sizes: 2D, 2D, and 1D, with 64, 64, and 32 filters, respectively. These convolutional layers play distinct roles in capturing spatial features and patterns within the EEG data.

The network architecture begins with a 2D convolutional layer, tasked with extracting spatial features from the input data. Subsequently, max pooling is applied to downsample the feature maps by selecting the maximum values from local regions. Dropout is incorporated as a regularization technique to mitigate overfitting during training.

Following the 2D CNN layer, we introduce a 2D convolutional layer, which further enhances the capture of spatial features. Max Pooling is again applied to the downsampled feature maps, and Dropout is employed for regularization.

This structure is then followed by a 1D convolutional layer, designed for feature extraction from sequences. Max pooling is again applied to the downsampled feature maps, with dropout used for regularization. This arrangement enhances the network's capability to effectively capture both spatial and sequential features.

The output from the CNN layers, consisting of sequences of feature maps, is reshaped into an appropriate format for the LSTM. This reshaping results in sequences with dimensions 32 features. These sequences are processed by 32 LSTM units, which play a pivotal role. These LSTM cells specialize in capturing and comprehending temporal dependencies within the EEG signals. As EEG data is inherently time-varying, with emotions evolving over time, LSTMs are exceptionally well-suited to recognize how emotional states evolve, making them indispensable for our model's success in emotion recognition.

This architectural synergy of CNN and LSTM layers empowers our model to process information effectively on two fronts. Firstly, the CNN layers focus on spatial relationships within EEG feature maps, adept at identifying relevant patterns and structures. Secondly, LSTM layers handle the temporal aspect, tracking how emotional states change across different time points. This combination provides a holistic approach to understanding emotions in EEG data.

The flexibility of this architecture makes it highly suitable for tasks where input data exhibits both spatial and temporal characteristics, as is the case in EEG signal analysis. It facilitates feature extraction at various scales and levels of abstraction, thereby enhancing the model's ability to recognize emotions accurately. To harness the full potential of this architecture, meticulous hyperparameter tuning and customized training procedures tailored to your specific dataset and task are essential. This ensures that the model is finely tuned to excel in recognizing emotions from EEG signals, making it a valuable tool in emotion analysis. Finally, our network incorporates a fully connected layer followed by an output layer. These layers are crucial for interpreting the learned features and producing the final classification results.

For classification and recognition, our model outputs results in four emotional dimensions: arousal, valence, dominance, and liking. To achieve this, we employ a classification algorithm using the mini-batch gradient optimization technique and a cross-entropy loss function. This combination allows our model to learn and fine-tune its parameters efficiently, enhancing its ability to recognize emotions in EEG data.

In summary, the fusion of CNN and LSTM architectures proves to be a highly efficient approach for EEG signal analysis. This synergy enables our model to process both spatial relationships within EEG feature maps and temporal relationships across different time points. Furthermore, when coupled with automated feature extraction methods like Auto-Encoder, our model can achieve the highest accuracy in emotion recognition.

Regarding the Rectified Linear Unit (ReLU) activation function, it serves as a pivotal component in our neural network layers. ReLU is preferred because it effectively activates only the relevant nodes in the network while keeping others off, making the training process computationally lighter. This characteristic is crucial, especially when dealing with large datasets like EEG signals. Moreover, ReLU's inherent sparsity in node activation helps reduce parameter interdependence, which, in turn, improves the model's ability to generalize well to unseen data, mitigating overfitting risks.

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL SETUP

The categorical cross-entropy loss function has been effectively utilized in an end-to-end training process for our model:

$$L(y, \hat{y}) = -\sum_{(C=1)}^{M} y_c log(\hat{y}_c) \qquad (1)$$

where M is the number of classes, in our case M is equal to 2 (low or high) and $\hat{y}_c$ is the predicted label, and y_c is the true label.

The deep neural network performs by minimizing the loss function. The stochastic gradient descent (SGD) optimizer has been used to train the proposed model with a learning rate equal to 0.01 and a decay equal to 0.82. The proposed model is validated using 10-fold cross validation, where the dataset is divided into 10-folds and each time, a fold is leaved out as a test data and the remaining folds are used to train the model. The hyper-parameters that influence the model performance are optimized and fixed by the grid search approach. Several optimization techniques have been applied and the SGD has illustrated the best accuracy. The performance of our model to classify the four emotions: arousal, valence, dominance and liking has been evaluated using four indexes: classification accuracy, precision and recall. Let define:

- Classification accuracy (ACC) is the proportion of correct predictions made by the classifier. It reads to the ratio between the number of true positives and true negatives to the total number of predictions made.
- Precision is the proportion of true positive predictions among all the positive predictions made. It is the number of true positives divided by the sum of the true positives and false positives.
- Recall is the proportion of true positive predictions among all the actual positive cases. It is the number of true positives divided by the sum of the true positives and false negatives.

B. MODEL VALIDATION

The proposed model has been trained on the four dimensions of the DEAP dataset. Figure 4 illustrates the evolution of training and validation loss (categorical cross-entropy loss function) of the classification of arousal emotion. As is evident in Figure 4, the proposed model reduces the training/validation loss and the validation loss converges quickly.

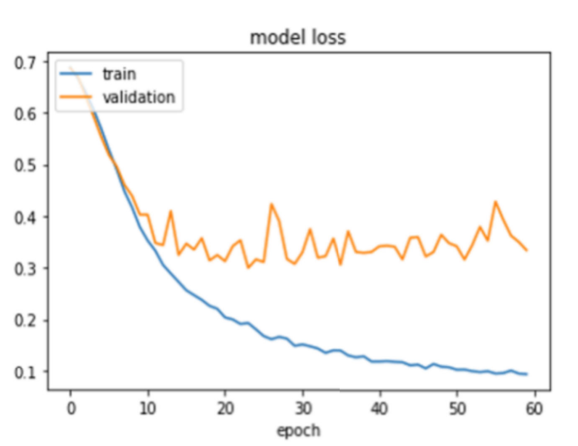


FIGURE 4. The evolution of the cross-entropy loss function for model training and validation to recognize the arousal emotion.

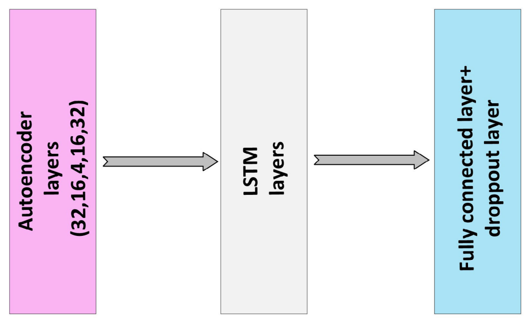


FIGURE 5. Ablation study: considered architecture where the CNN layers are removed (Experience 1).

In order to evaluate the performance of our framework on the four emotional dimensions: arousal, valence, dominance and liking, two experiments have been conducted for an ablation study. This helps us to get a complete understanding of how well the proposed model is able to classify different emotions and to highlight the impact of each component of the network. The first experiment aims to illustrate the positive impact of the AE-LSTM combination, where the CNN layers are removed. Therefore, the network of the first experiment is composed of the autoencoder layers combined with the 32 LSTM units. Then a fully connected layer and an output layer followed (see Figure 5). The second ablation experiment includes the CNN-LSTM combination to compare performance with the complete architecture. (see Figure 5).

The purpose of the second experiment is to test the efficiency of autoencoder, and if the combined CNN-LSTM network achieve the same results without the use of autoencoder to extract pertinent features. The architecture of the network is illustrated in Figure 6.

The third experiment tests the efficiency of LSTM by removing it from the proposed architecture.

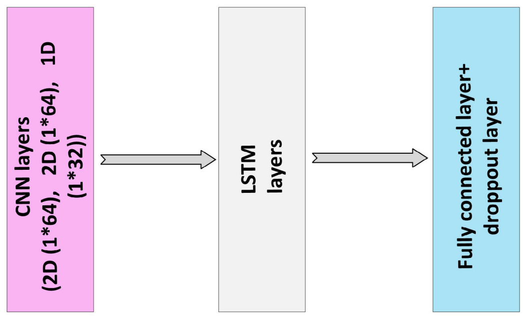


FIGURE 6. Ablation study: considered architecture where the autoencoder layers are removed (Experience 2).

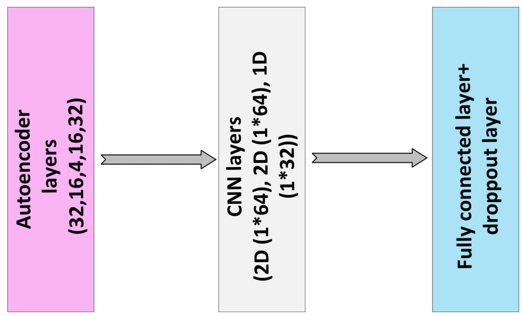


FIGURE 7. Ablation study: considered architecture where the LSTM layers are removed (Experience 3).

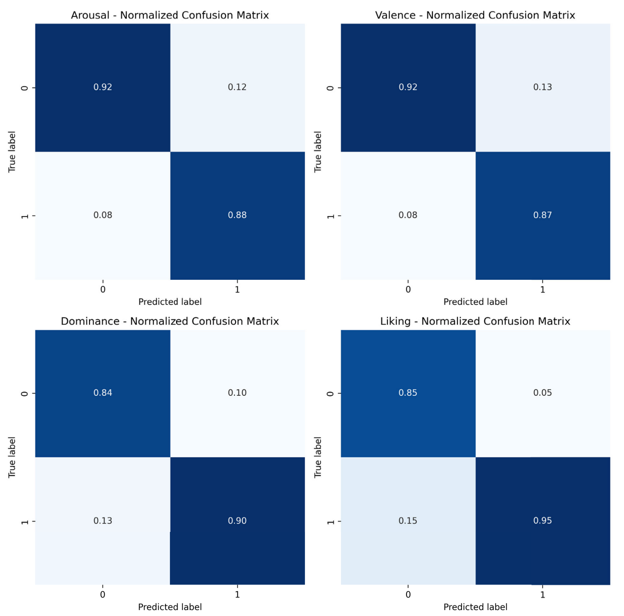


FIGURE 8. Normalized Confusion Matrices for Emotion Recognition: Arousal, Valence, Dominance, and Liking.

Tables 1 illustrates the comparison between the proposed method for emotion recognition based on EEG signal and the networks proposed in the conducted experiments (Figure 5, 6 and 8). The comparison illustrates the effectiveness of the proposed approach against the other networks. The proposed model was able to detect the four emotions with accuracy higher than 80%, the best classification accuracy was denoted by the liking emotion with 90.84% and arousal emotion 90.04%. The valence and dominance emotion recognition accuracies were 89.97% and 87.73% respectively.

The comparison between the results of all networks illustrates the importance of using the three components. The use of AE as feature extraction method applied before the CNN-LSTM combination improved the results of all performance measures (ACC, precision and recall), except for the recall index for the arousal emotion recognition, a minor degradation can be pointed (0.46%), but it can be justified by the variation of the 10-folds results (std is smaller when AE is used).

The impact of the CNN is very remarkable, the CNN has significantly improved the results of all emotions recognition with very considerable differences, for instance, the classification accuracy of the arousal, valence dominance and liking has been improved by 7.06%, 6.85%, 5.16% and 2.6% respectively.

The LSTM component plays a crucial role in emotion recognition by capturing temporal dependencies in the EEG signals. Our ablation study (experiment 3) showed that when the LSTM layer was removed, there was a significant drop in performance across all metrics. Specifically, the absence of LSTM resulted in lower accuracy and precision for all emotion categories. This underscores the importance of LSTM in modeling the sequential nature of EEG data, allowing the network to better understand the temporal dynamics associated with different emotional states.

The confusion matrices reveal that the model generally performs well across the four emotional dimensions. The normalized values show that the model has high accuracy in correctly identifying instances of each emotion. The model recognizes the "Liking" emotion most accurately, as indicated by the higher diagonal values. On the other hand, the "Dominance" emotion is recognized with slightly less accuracy.

C. COMPARISON WITH RELATED WORKS

Table 2 summarizes the results achieved in literature for emotion recognition it also compares the proposed method with reported works. The compared methods share the same conditions as our model, they have been evaluated on the DEAP dataset with binary classification of the emotions. In this study, we have applied our model on the four emotions, whereas the other studies have applied it only on 2 dimensions. Table 2 illustrates the superiority of our model compared to other studies results. It can recognize the arousal and valence emotions with accuracies equal to 90.04% and 89.97%. With these results, it exceeds the results presented on [39] by 4.39%, 2.52% and 2.85% for the recognition of arousal, valence and liking emotions respectively. To the best of our knweledge, the work reported in [39] is considered as

TABLE 1. Comparison of ablation experiments classification results.

Emotions	Arousal%			Valence%			Dominance%			Liking%		
Evaluation indexes	ACC	Precision	Recall	ACC	Precision	Recall	ACC	Precision	Recall	ACC	Precision	Recall
Experiment 1: AE+LSTM	82.98 ±0.21	79.13 ±0.34	84.09 ±0.25	83.12 ±0.08	80.34 ±0.18	78.51 ±0.21	82.57 ±0.62	73.34 ±0.21	72.34 ±0.01	88.24 ±0.68	88.45 ±0.54	86.19 ±0.48
Experiment 2: CNN+LSTM	84.02 ±0.67	82.43 ±0.21	84.54 ±0.34	85.67 ±0.01	82.68 ±0.53	78.54 ±0.67	81.34 ±0.53	73.23 ±0.08	72.47 ±0.25	87.98 ±0.07	87.92 ±0.32	87.23 ±0.81
Experiment 3: AE+CNN	85.92 ±0.5	82.76 ±0.7	83.72 ±0.32	84.89 ±0.2	81.93 ±0.12	80.81 ±0.38	84.32 ±0.34	72.98 ±0.02	74.67 ±0.43	86.72 ±0.02	85.21 ±0.13	87.83 ±0.23
Proposed model: AE+CNN+LSTM	90.04 ±0.55	84.34 ±0.02	84.08 ±0.07	89.97 ±0.31	85.23 ±0.08	79.73 ±0.07	87.73 ±0.18	73.34 ±0.05	76.32 ±0.25	90.84 ±0.01	90.02 ±0.61	88.84 ±0.53

TABLE 2. Comparison of our auto-encoder+ CNN-LSTM model and previous emotion recognition studies.

Methods	Arousal	Valence	Dominance	Liking
[29]	60.72%	62.4%	-	-
[50]	73.125%	72.1%	-	-
[38]	73.36%	81.4%	-	-
[39]	85.65%	87.45%	-	87.99%
[40]	81.43%	76.70%	-	-
[51]	76.30%	76.54%	-	-
[52]	64.33%	64.25%	-	-
[47]	92.86%	89.49%	-	-
[53]	90.01%	90.01%	-	-
Proposed method	90.04%	89.97%	87.73%	90,84%

the best until now and the only one which applied the same recognition model on 3 emotion dimensions.

D. DISCUSSION

The proposed emotion recognition model is based on a combination of different deep-learning layers, which can be categorized by their role in two components: (1) Autoencoder for feature extraction from raw EEG signal and (2) classification component which is a combination of 2D-2D-1D-CNN and LSTM layers. The proposed method has shown its efficiency compared to other state-of-the-art methods.

The CNN-SAE-DNN model reported in [47], has achieved 92.86% accuracy for "arousal" and 89.49% for "valence". While our AE-CNN-LSTM model achieved 90.04% for "arousal" and 89.97% for "valence", the differences in performance may be due to variations in preprocessing and windowing techniques. Reference [47] have employed different filtering and segmentation methods, whereas our study used raw data with minimal preprocessing. These methodological differences can significantly influence the outcomes and highlight the importance of standardized pre-processing steps for fair comparisons. The proposed method outperforms previous studies, such as [38], where only LSTM was used to classify emotions. The latter considers only temporal dependencies, which may explain why the use of LSTM layers alone cannot fully recognize emotions. In contrast, our method combines spatial and temporal feature extraction through CNN and LSTM layers, providing a more comprehensive approach to emotion recognition.

The contribution of each component of our architecture has been illustrated by an ablation study. The first point derived from these comparisons is the effectiveness of the autoencoder for feature extraction has been confirmed Table 1. This is not a surprising result, because we are dealing with raw signals, so a feature extraction is a required step in the learning process.

The second comment regarding Table 2 is the importance of the CNN layers and how they have significantly improved the results. In [40], authors have also used a CNN-LSTM model but they used only a 1D CNN. This supports the conclusion of [49] about the strength of 2D-CNN when spatial information is combined with either spectral or temporal information. The third note concerns the model generalization on four emotion dimensions. As we applied the model on the four emotions considered on the DEAP dataset, it is important to discuss the model generalization. It is clear that the proposed model outperforms the results of other studies on all dimensions, but we have noted that the results obtained on the arousal and liking dimensions were the most superior. If we refer to the two emotions definitions, we can find a correlation between these emotions which can justify the obtained results. In fact, liking is an emotion that refers to the positive feelings, which is often accompanied by a desire to interact with the object of our liking. Arousal, on the other hand, is a physiological response that refers to the level of activity or excitement in the body. The experience of positive emotions like liking, involves body excitation, as evidenced by an increase in heart rate, blood pressure, and other physiological responses. However, it is important to note that liking and arousal are two separate and distinct emotions, and they can be experienced independently of each other.

While the proposed framework demonstrates strong performance on the DEAP dataset, future work will focus on validating its effectiveness on other EEG datasets to further establish its generalizability. Additionally, evaluating the framework in real-time applications, such as emotion-aware systems and adaptive human-computer interaction, represents an important next step to assess its practical deployment in dynamic environments. These directions will build on the solid foundation established by this study and further enhance the applicability of the proposed method.

V. CONCLUSION

We proposed in this paper a deep learning based model for emotion recognition based on EEG signal analysis. The first component of the model is dedicated to the decomposition of signals and the extraction of features that will support the classifier in discriminating four emotions: arousal, valence, dominance, and liking, which are the labels of the collected EEG signals composing the DEAP dataset. A dense autoencoder model has been proposed for this step. The use of AE has been validated in results section, where the classifier has been tested with and without it, and the results have confirmed its efficiency. Then, the extracted features were injected to a combined CNN-LSTM network in order to learn and identify patterns in the data that are indicative of different emotions. The combination of the 2D-CNN and LSTM layers in a single classifier generates a classification model able to analyze both the spatial and temporal relationships in the EEG data in order to improve the recognition accuracy. The comparative results demonstrated the effectiveness of our method, it achieved classification accuracies equal to 90.04%, 89.97%, 87.73% and 90.84%for arousal, valence, dominance and liking emotion recognition. For future works, we plan the used of the classification model for developing a real-word emotion recognition framework which can be used to recognize different psychological disorders.

DATA AVAILABILITY

The DEAP dataset has been used in this research, it can accessed publicly: https://www.eecs.qmul.ac.uk/mmv/datasets/deap/

REFERENCES

[1] P. Yang, "Prediction model of paroxysmal atrial fibrillation based on pattern recognition and ensemble CNN-LSTM," *Zhejiang Daxue Xuebao*, vol. 54, no. 1, pp. 1039–1048, 2020.

[2] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol*, vol. 29, pp. 306–310, Jun. 1970.

[3] Z. Yin, "Cross-subject EEG feature selection for emotion recognition using transfer recursive feature elimination," *Frontiers Neurorobot*, vol. 11, p. 19, 2017, doi: 10.3389/FNBOT.2017.00019.

[4] I. Jemal, N. Mezghani, L. Abou-Abbas, and A. Mitiche, "An interpretable deep learning classifier for epileptic seizure prediction using EEG data," *IEEE Access*, vol. 10, pp. 60141–60150, 2022, doi: 10.1109/ACCESS.2022.3176367.

[5] L. Abou-Abbas, K. Henni, I. Jemal, A. Mitiche, and N. Mezghani, "Patient-independent epileptic seizure detection by stable feature selection," *Expert Syst. Appl.*, vol. 232, Dec. 2023, Art. no. 120585, doi: 10.1016/j.eswa.2023.120585.

[6] L. Abou-Abbas, I. Jemal, K. Henni, Y. Ouakrim, A. Mitiche, and N. Mezghani, "EEG oscillatory power and complexity for epileptic seizure detection," *Appl. Sci.*, vol. 12, no. 9, p. 4181, Apr. 2022, doi: 10.3390/app12094181.

[7] I. Jemal, A. Mitiche, and N. Mezghani, "A study of EEG feature complexity in epileptic seizure prediction," *Appl. Sci.*, vol. 11, no. 4, p. 1579, Feb. 2021, doi: 10.3390/app11041579.

[8] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, 2014, doi: 10.1561/2000000039.

[9] X. Yang, W. Deng, and J. Yao, "Neural adaptive dynamic surface asymptotic tracking control of hydraulic manipulators with guaranteed transient performance," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 1–11, May 2022, doi: 10.1109/TNNLS.2022.3141463.

[10] X. Yang, W. Deng, and J. Yao, "Neural network based output feedback control for DC motors with asymptotic stability," *Mech. Syst. Signal Process.*, vol. 164, Feb. 2022, Art. no. 108288, doi: 10.1016/j.ymssp.2021.108288.

[11] Y.-H. Kwon, S.-B. Shin, and S.-D. Kim, "Electroencephalography based fusion two-dimensional (2D)-convolution neural networks (CNN) model for emotion recognition system," *Sensors*, vol. 18, no. 5, p. 1383, Apr. 2018, doi: 10.3390/s18051383.

[12] Y.-J. Liu, M. Yu, G. Zhao, J. Song, Y. Ge, and Y. Shi, "Real-time movie-induced discrete emotion recognition from EEG signals," *IEEE Trans. Affect. Comput.*, vol. 9, no. 4, pp. 550–562, Oct. 2018, doi: 10.1109/TAFFC.2017.2660485.

[13] E. Akleman, "Deep learning," *Computer*, vol. 53, no. 9, pp. 1–17, Sep. 2020.

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[15] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update," *J. Neural Eng.*, vol. 15, no. 3, Jun. 2018, Art. no. 031005.

[16] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2014, pp. 3104–3112.

[17] W. L. Zheng, "EEG-based emotion classification using deep belief networks," in *Proc. IEEE Int. Conf. Multimedia Expo.*, May 2014, pp. 1–6, doi: 10.1109/ICME.2014.6890166.

[18] S. Paul, A. Mazumder, P. Ghosh, D. N. Tibarewala, and G. Vimalarani, "EEG based emotion recognition system using MFDFA as feature extractor," in *Proc. Int. Conf. Robot., Autom., Control Embedded Syst. (RACE)*, Feb. 2015, pp. 1–5.

[19] P. Ackermann, "EEG-based automatic emotion recognition: Feature extraction, selection and classification methods E-Health networking, applications and services (Healthcom)," in *Proc. IEEE 18th Int. Conf.*, Jun. 2016, pp. 1–6.

[20] J. Qiu, "Great strides of China's space programmes," *Nat. Sci. Rev.*, vol. 4, no. 2, pp. 264–268, Mar. 2017.

[21] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from EEG signals using multidimensional information in EMD domain," *BioMed Res. Int.*, vol. 2017, pp. 1–9, Apr. 2017.

[22] A. Mert and A. Akan, "Emotion recognition from EEG signals by using multivariate empirical mode decomposition," *Pattern Anal. Appl.*, vol. 21, no. 1, pp. 81–89, Feb. 2018.

[23] S. Guzel Aydin, T. Kaya, and H. Guler, "Wavelet-based study of valence–arousal model of emotions on EEG signals with LabVIEW," *Brain Informat.*, vol. 3, no. 2, pp. 109–117, Jun. 2016.

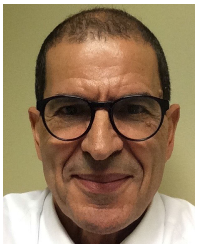
[24] P. Pandey and K. R. Seeja, "Emotional state recognition with EEG signals using subject independent approach," in *Lecture Notes on Data Engineering and Communications Technologies*. Cham, Switzerland: Springer, 2019, pp. 117–124.

[25] P. Pandey, "Subject-independent emotion detection from EEG signals using deep neural network," in *Proc. Int. Conf. Innov. Comput. Commun.*, 2019, pp. 1–20.
[26] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using EEG signals: A survey," *IEEE Trans. Affective Comput.*, vol. 2, no. 1, pp. 1–20, May 2017.
[27] Y. H. Liu, "Single-trial EEG-based emotion recognition using kernel eigen-emotion pattern and adaptive support vector machine," in *Proc. 35th Annu. Int. Conf.*, 2013, pp. 4306–4309.
[28] Z. Lan, O. Sourina, L. Wang, and Y. Liu, "Real-time EEG-based emotion monitoring using stable features," *Vis. Comput.*, vol. 32, no. 3, pp. 347–358, Mar. 2016.
[29] J. Atkinson and D. Campos, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," *Expert Syst. Appl.*, vol. 47, pp. 35–41, Apr. 2016.
[30] Z. Mohammadi, J. Frounchi, and M. Amiri, "Wavelet-based emotion recognition system using EEG signal," *Neural Comput. Appl.*, vol. 28, no. 8, pp. 1985–1990, Aug. 2017.
[31] Y. Jiang, "A novel distributed multitask fuzzy clustering algorithm for automatic MR brain image segmentation," *J. Med. Syst.*, vol. 43, p. 118, May 2019, doi: 10.1007/s10916-019-1245-1.
[32] W.-L. Zheng and B. L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Trans. Auton. Mental Dev.*, vol. 7, no. 1, pp. 162–175, May 2015, doi: 10.1109/TAMD.2015.2431497.
[33] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis ;Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012, doi: 10.1109/T-AFFC.2011.15.
[34] F. Wang, "Data augmentation for EEG-based emotion recognition with deep convolutional neural networks," in *Proc. Int. Conf. Multimedia Model.*, 2018, pp. 82–93.
[35] H. Zeng, Z. Wu, J. Zhang, C. Yang, H. Zhang, G. Dai, and W. Kong, "EEG emotion classification using an improved SincNet-based deep learning model," *Brain Sci.*, vol. 9, no. 11, p. 326, Nov. 2019.
[36] A. Garg, A. Kapoor, A. K. Bedi, and R. K. Sunkaria, "Merged LSTM model for emotion classification using EEG signals," in *Proc. Int. Conf. Data Sci. Eng. (ICDSE)*, Sep. 2019, pp. 139–143.
[37] H. Sunhee, "Learning CNN features from DE features for EEG based emotion recognition," *Pattern Anal. Appl.*, vol. 2, no. 1, pp. 1–13, Apr. 2019.
[38] T. Samarth, "Sing deep and convolutional neural networks for accurate emotion classification on DEAP dataset," in *Proc. 29th IAAI Conf.*, 2017, pp. 1–12.
[39] S. Alhagry and A. Aly, "Emotion recognition based on EEG using LSTM recurrent neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 10, pp. 355–358, 2017.
[40] Q. Li, Y. Liu, Y. Shang, Q. Zhang, and F. Yan, "Deep sparse autoencoder and recursive neural network for EEG emotion recognition," *Entropy*, vol. 24, no. 9, p. 1187, Aug. 2022, doi: 10.3390/e24091187.
[41] K.-Y. Wang, Y.-L. Ho, Y.-D. Huang, and W.-C. Fang, "Design of intelligent EEG system for human emotion recognition with convolutional neural network," in *Proc. IEEE Int. Conf. Artif. Intell. Circuits Syst. (AICAS)*, Mar. 2019, pp. 142–145.
[42] P. J. Lang, "The emotion probe: Studies of motivation and attention," *Amer. Psychologist*, vol. 50, no. 5, pp. 372–385, 1995.
[43] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time EEG-based human emotion recognition and visualization," in *Proc. Int. Conf. Cyberworlds*, Oct. 2010, pp. 262–269, doi: 10.1109/CW.2010.37.
[44] J. N. Acharya, A. Hani, J. Cheek, P. Thirumala, and T. N. Tsuchida, "American clinical neurophysiology society guideline 2: Guidelines for standard electrode position nomenclature," *J. Clin. Neurophysiology*, vol. 33, no. 4, pp. 308–311, 2016.
[45] S. Petridis, B. Martinez, and M. Pantic, "The MAHNOB laughter database," *Image Vis. Comput.*, vol. 31, no. 2, pp. 186–202, Feb. 2013.
[46] J. D. Morris, "Observations: SAM: The self-assessment manikin; An efficient cross-cultural measurement of emotional response," *J. Advertising Res.*, vol. 35, no. 6, pp. 63–68, 1995.
[47] P. Pandey and K. R. Seeja, "Subject independent emotion recognition from EEG using VMD and deep learning," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 5, pp. 1730–1738, May 2022.
[48] J. Liu, G. Wu, Y. Luo, S. Qiu, S. Yang, W. Li, and Y. Bi, "EEG-based emotion classification using a deep neural network and sparse autoencoder," *Frontiers Syst. Neurosci.*, vol. 14, pp. 1–43, Sep. 2020, doi: 10.3389/fnsys.2020.00043.
[49] D. Shah, G. Gopan, and N. Sinha, "An investigation of the multi-dimensional (1D vs. 2D vs. 3D) analyses of EEG signals using traditional methods and deep learning-based methods," *Frontiers Signal Process.*, vol. 2, pp. 1–29, Jul. 2022.
[50] Y. Wang, Z. Huang, B. McCane, and P. Neo, "EmotioNet: A 3-D convolutional neural network for EEG-based emotion recognition," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–7.
[51] A. Topic and M. Russo, "Emotion recognition based on EEG feature maps through deep learning network," *Eng. Sci. Technol., Int. J.*, vol. 24, no. 6, pp. 1442–1454, Dec. 2021.
[52] R. Li, C. Ren, X. Zhang, and B. Hu, "A novel ensemble learning method using multiple objective particle swarm optimization for subject-independent EEG-based emotion recognition," *Comput. Biol. Med.*, vol. 140, Jan. 2022, Art. no. 105080.
[53] M. A. H. Akhand, M. A. Maria, M. A. S. Kamal, and K. Murase, "Improved EEG-based emotion recognition through information enhancement in connectivity feature map," *Sci. Rep.*, vol. 13, no. 1, Aug. 2023, Art. no. 13804.

KHADIDJA HENNI received the master's degree in computer science and the Ph.D. degree from the University of Science and Technology of Oran, in 2017.

Following the Ph.D. degree, she conducted postdoctoral research with TÉLUQ University and the LIO Laboratory, CR-CHUM, Montreal, Canada, where she specialized further in machine learning and artificial intelligence. She is currently a Professor with the Department of Science and Technology, TÉLUQ, and a Researcher with the Institute of Applied Artificial Intelligence. Throughout her career, she has made significant contributions to the development of various clustering and feature selection algorithms. Her research interests include applying artificial intelligence techniques to biomedical and bioinformatics research.

NEILA MEZGHANI received the degree in telecommunications engineering from the Higher School of Telecommunications of Tunis (Sup'Com), the master's degree in information technology from the National School of Engineers of Tunis, and the Ph.D. degree from the Institut National de Recherche Scientifique—Centre Énergie Matériaux et Télécommunications, Montréal. She is currently a Data Scientist Professor with Université TÉLUQ (Quebec University) and a Researcher with the Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CR-CHUM). She is the author of two patents and about 100 peer-reviewed publications in renowned scientific journals and international conferences. Her research interests include biomedical data mining and classification, artificial intelligence, decision support systems in the medical field, and mobile health. She was the Canada research chair in biomedical data mining.



AMAR MITICHE received the Licence ès Sciences degree in mathematics from the University of Algiers and the Ph.D. degree in computer science from The University of Texas at Austin. He is currently a Professor with the Department of Telecommunications (INRS-EMT), Institut National de la Recherche Scientifique (INRS), Montreal, QC, Canada. His research interests include computer vision and pattern recognition. He has written several articles on the subjects and three books: *Computational Analysis of Visual Motion* (Plenum Press, 1994), *Variational and Level Set Methods in Image Segmentation* (Springer, 2011), with Ismail Ben Ayed, and *Computer Vision Analysis of Image Motion by Variational Methods* (Springer, 2014), with J. K. Aggarwal. His current research interests include image segmentation, image motion analysis, and pattern classification by neural networks.

LINA ABOU-ABBAS received the Ph.D. degree in electrical engineering from the Ecole de Technologie Superieure, Montreal, QC, Canada, in December 2016. After the Ph.D. degree and until May 2020, she was a Postdoctoral Fellow with the Faculty of Medicine, Department of Neurology and Neurosurgery, McGill University. Her position was cross-appointed with the Douglas University Mental Health Institute and the McGill University Health Center. She contributed to the longitudinal behavioral data collection and organization by designing a web-based data capturing and archiving platform that allows multisite collaboration (Halifax, McMaster, and McGill Universities). In August 2020, she joined as a Researcher with the LICEF Institute and the Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CR-CHUM). Her research interest includes understanding the brain basis of behavioral disorders. Her current research interests include machine learning, deep learning, pattern recognition, and signal processing. She was a Core Member of the Transforming Autism Care Consortium and Canadian Autism Neuroinformatics platform funded by Brain Canada and grouping a national and international multi-disciplinary researcher. She was awarded Quebec Autism Research Training (QART) Program Fellowship, in 2019. She was a recipient of the Fonds du Quebec—Nature et Technologie Postdoctoral Fellowship, from 2019 to 2022.

AMEL BENAZZA-BEN YAHIA received the Engineering degree from the National Institute of Telecommunications, Evry, France, in 1988, the Ph.D. degree from the University Paris-Sud (XI), Paris, France, in 1993, and the Habilitation Universitaire degree from the Ecole Supérieure des Communications (SUP'COM), Tunis, Tunisia, in 2003. She is currently a Full Professor with the Department of Applied Mathematics, Signal Processing and Communications, SUP'COM, University of Carthage, Tunisia. She is also a Research Scientist with the Communication, Signal and Image Laboratory (COSIM). Her research interests include image compression, signal denoising, and content-based image retrieval.